

AD _____

Award Number: DAMD17-01-1-0395

TITLE: Development, Optimization and Evaluation of CAD System
for Breast Cancer

PRINCIPAL INVESTIGATOR: Wei Qian, Ph.D.

CONTRACTING ORGANIZATION: University of South Florida
Tampa, FL 33620-7900

REPORT DATE: June 2003

TYPE OF REPORT: Annual Summary

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

DISTRIBUTION STATEMENT: Approved for Public Release;
Distribution Unlimited

The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy or decision unless so designated by other documentation.

20030902 127

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 074-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing this collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503

1. AGENCY USE ONLY (Leave blank)		2. REPORT DATE June 2003	3. REPORT TYPE AND DATES COVERED Annual Summary (14 May 2002 - 14 May 2003)	
4. TITLE AND SUBTITLE Development, Optimization and Evaluation of CAD System for Breast Cancer			5. FUNDING NUMBERS DAMD17-01-1-0395	
6. AUTHOR(S) Wei Qian, Ph.D.				
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) University of South Florida Tampa, FL 33620-7900 E-Mail: qianw@ Moffitt.usf.edu			8. PERFORMING ORGANIZATION REPORT NUMBER	
9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES) U.S. Army Medical Research and Materiel Command Fort Detrick, Maryland 21702-5012			10. SPONSORING / MONITORING AGENCY REPORT NUMBER	
11. SUPPLEMENTARY NOTES				
12a. DISTRIBUTION / AVAILABILITY STATEMENT Approved for Public Release; Distribution Unlimited				12b. DISTRIBUTION CODE
13. ABSTRACT (Maximum 200 Words) <p>A 12 weeks summer research training program for undergraduate students focused on understanding of digital mammography, medical imaging and breast cancers. Recruitment of program participants was recruited through poster distribution to more than 100 college campuses. Selection of program participants applications were evaluated based on</p> <ol style="list-style-type: none">1. letter of recommendation2. transcript of grades3. 500-word essay stating why you want to attend Moffitt's Summer training Program. <p>113 students applied and 46 of them completed the above 3 steps and finally only 6 of the students have been accepted for the training program each year.</p> <p>The purpose was to teach trainees to have basic theory and technique on the breast cancer study. Trainees practiced on development of CAD modules and completed small scale projects related to CAD modules for breast cancer under mentors' direction. They submitted scientific report before the end of training program.</p>				
14. SUBJECT TERMS Digital mammography, Computer-aided-diagnosis, Database, Image processing, Evaluation				15. NUMBER OF PAGES 9
				16. PRICE CODE
17. SECURITY CLASSIFICATION OF REPORT Unclassified	18. SECURITY CLASSIFICATION OF THIS PAGE Unclassified	19. SECURITY CLASSIFICATION OF ABSTRACT Unclassified	20. LIMITATION OF ABSTRACT Unlimited	

NSN 7540-01-280-5500

Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std. Z39-18
298-102

Table of Contents

Cover.....	1
SF 298.....	2
Table of Contents.....	3
Introduction.....	4
Body.....	5
Key Research Accomplishments.....	8
Reportable Outcomes.....	8
Conclusions.....	9
References.....	
Appendices.....	

Development, Optimization and Evaluation of CAD System for Breast Cancer

Introduction

A 12 weeks summer research training program for undergraduate students focused on understanding of digital mammography, medical imaging and breast cancers. Recruitment of program participants was recruited through poster distribution to more than 100 college campuses. Selection of program participants applications were evaluated based on

1. letter of recommendation
2. transcript of grades
3. 500-word essay stating why you want to attend Moffitt's Summer training Program.

113 students applied and 46 of them completed the above 3 steps and finally only 6 of the students have been accepted for the training program each year. In order to attract and train undergraduates who have interests and potentially become excellent researcher on the breast cancer study, we designed this undergraduate training program. This proposed training program is planned for undergraduates majoring in electrical engineering, medicine science or radiology. The program was designed as following: They were guided to visit the clinical site of related devices for mammography, the biopsy process for patients with suspicious breast tumor in Moffitt cancer center. They visited the process of film-screen image converting and transmitting and display of processed medical images in the laboratory of Digital Medical Imaging Program (DMIP) in the cancer research institute. It helped trainees to set up perceptive understanding on mammography for breast cancer diagnosis. The trainees were given a series of lectures on the basic principle of mammography, medical imaging, image processing, CAD methods, generation of databases and truth files, evaluation of CAD methods by means of lecture and seminars. The purpose was to teach trainees to have basic theory and technique on the breast cancer study. Trainees practiced on development of CAD modules and completed small-scale projects related to CAD modules for breast cancer under mentors' direction. They submitted scientific report before the end of training program.

BODY OF REPORT

The following activities were organized for the training program

Orientation

1. Welcome dinner
2. Lab. Safety and Hazardous Materials management workshop
3. Tour of Library with science reference librarians
4. Introduction and Schedule of Lab. seminars, topics included

Presentation Skill Training (How to design, prepare and present a research work)

1. Writing a scientific abstract
2. Poster format Presentation
3. Oral format presentation

Career Opportunity Introduction (*Information on applying to graduate and medical programs*)

1. The road to Graduate School
2. Career paths after Graduate School
3. The road to Medical School
4. Research in Medical School
5. Career paths after Medical School

The training program is designed as four stages: visiting, lecturing, practicing and doing projects.

(1). Visiting

In order to get a basic understanding of mammography and the role of medical image in the breast cancer diagnosis, the trainees will be scheduled to visit the Moffitt cancer center, where they will see the clinical implementation of X-ray mammography or direct digital mammography devices to learn the process of mammography. In addition, they will observe the biopsy method for the patient who is diagnosed with suspicious tumor based on their images. Noticing the pain that the patients suffer from during the biopsy, the trainees will get a deep impression about how important correct diagnosis of breast cancer based on the mammography for patients is. The undergraduates will visit the laboratory of DMIP to learn the conversion process of the film-screening image into a digital one, the generation of medical image database, the display of medical image on computer, and medical image processing. It will give trainees the understanding that CAD methods for mass and MCCs detection are helpful for breast cancer diagnosis.

(2). Lecturing

A series of lectures or seminars by mentors and other scientists in breast cancer study are planned for training undergraduates for them to obtain thorough understanding of mammography and find more interest in the study of CAD for breast cancer, which includes:

- (a) General introduction of the development of mammography, emphasizing on its importance in the future and on the diagnosis of breast cancer.

- (b) The features of breast tumor, the behaviors of mass and microcalcification clusters (MCCs) on X-ray images, the difference between benign and malignant tumor.
- (c) The generation of image database and truth files.
- (d) Basic principle of medical image, including film-screen mammography and direct digital mammography.
- (e) Basic theory of image processing, including: enhancement, segmentation, feature extraction, selection and classification.
- (f) Basic principle of pattern recognition and feature classification.
- (g) The development of CAD methods for breast cancer detection. The emphasis will be placed on the detection of mass and MCCs with this technology.

(3). Practicing

In this training stage, undergraduates will be divided into three groups with different subjects based on their major and their interests. Students majoring in Electrical Engineering or Computer Science will mainly be divided into two groups: one group for development and optimization of CAD methods for mass/MCCs detection; another one for evaluation of CAD system with retrospective study. Trainees in medical science or radiology including some students in computer science will create one group to generate image databases and truth files.

(4). Projects

Mentors design a series of small-scale projects for trainees. These projects are related to the mentors' current research topics. Students under supervision of relevant mentor can do these projects or help mentors to do something for the projects. For example, the project (1) is a new project, this project proposes a novel computer-aided pathological diagnosis (CAPD) system for classifying the histo-differences between normal cells, cancer cells and subtypes of cancer cells, which is an innovative and efficient way for breast cancer early diagnosis. The student under this project helped a lot on breast cancer cell image collection and basic algorithm evaluation, which was very closely worked with mentor.

(1). Accurate histo-morphological distinction of breast cancer cells and distinction of specific subtypes of cancer cells are critical for optimum patient care and often can present a diagnostic challenge to a pathologist. This project proposes a novel computer-aided pathological diagnosis (CAPD) system for classifying the histo-differences between normal cells, cancer cells and subtypes of cancer cells, which is an innovative and efficient way for breast cancer early diagnosis. Molecular/Cellular-based approach to early detection is a revolution in diagnosis. This proposal will design a new CAPD system for detection and diagnosis of breast cancer with cellular images. It includes the following novel modules: (1) a novel adaptive fragmentary window filtering (AFWF) algorithm for circularity enhancement which is an important feature to identify normal and cancer cells, (2) modification of tree-structured nonlinear filtering (TSF), directional wavelet transform (DWT), tree-structured wavelet transform (TSWT), segmentation, feature extraction/selection, and classification modules. (3) Declustering for isolation of cell and nuclear. The CAPD method will be as a leading technology that can be exploited to address the current issues associated with using cellular image for early stage detection of breast cancer.

(2). Investigate the development of image preprocessing modules: It is the first and important stage in image processing. Several methods for this module have been developed with different algorithms, some of them the student has studied. We are trying to develop new adaptive image preprocessing modules. The trainees will take part in parts of this research project to learn about the development of contemporary CAD preprocessing methods. They performed comparison test among current developed methods through computer programming for different image databases. After finishing the work, the trainees wrote a report about the performance comparison of the CAD module with different image databases. During training, the trainees were encouraged to propose their own methods for this module and to do related computer programming.

(3). Generation of image databases and truth files: databases of film-screening mammography have been configured and are being expanded, the databases for direct digital mammography have been developed. The trainees generated databases for different kinds of images under the direction of mentors and research assistants. They learnt to construct the truth files for image databases. The student was required to submit reports about the constructed image database and related truth files.

(4). Development of Adaptive CAD module for false positive (FP) reduction: Sensitivity and false positive rate are two factors that greatly affect the clinical trial of CAD modules. For a long time, these two performance factors of CAD system were not suitable for the clinical use. Efforts are placed on searching an ideal method that can obtain high sensitivity and keep low false positive rate. The trainees worked on the investigation of the current literature. In our medical imaging Lab., a new kind of adaptive CAD system for false positive reduction has been developed. The trainees assisted the mentor to test the performance of developed CAD methods. The students submitted a report on the performance of new module for FP reduction.

(5). Evaluation of CAD system for breast cancer study: so many CAD systems for mass and MCCs detection have been developed, which need retrospective study and clinical analysis. It is suitable to do this evaluation study. Following the guidance of mentors, the trainee performed retrospective analysis on current CAD modules using the data sets developed through the Department of Defense Breast Cancer Research Program (BCRP) grant to the University of South Florida (USF) ([http://marathon.csee.usf.edu/Mammography/ Database.html](http://marathon.csee.usf.edu/Mammography/Database.html)) for film screening mammography or the databases developed by themselves for digital mammography. Being familiar with evaluation method for retrospective study, the students presented evaluation reports for different CAD systems.

(6). The research project involves three parts. *The first part* is for initial optimization of each module performed using standard signal processing criteria, analysis of simulated images and comparison of segmented images of mass area to ground truth files. The adaptive techniques are used to improve image preprocessing CAD modules. *The second part* is focused on a novel, fully automatic and highly efficient method for CAD system full optimization based on the clinical objective. The objective function is built from a set of 2000 case mammograms that will contain: 200 mammograms with no lesion (normals), 1800 mammograms with masses of irregular, circumscribed, microlobulated, obscured, ill-defined and spiculated as defined in BI-RADS. *In the third part*, is to design a statistic test to validate the optimized algorithm model. This statistical method is the Hypothesis Testing with the Null Hypothesis (NH) of that the two parameter settings of the CAD system have the same performance. The students submitted a report on the above study.

Assessment of the Program and Evaluation of the Students

The assessment of the program is based on whether the smooth implementation of the whole program structure, starting from recruitment of the participants, until to the end of the final project report. The participation of the mentors/scientists included, Wei Qian, Ph.D. (Associate Professor), John Heine, Ph.D. (Assistant Professor), Robert Clark M.D. (Professor), Xuejun Sun, Ph.D. (Research Associate). All participant students are good at following rules. They worked hard, but, they are "too young" for doing advanced projects. We don't want to design small projects that are only for their practice and not useful for our research goal. All the projects assigned to them are parts of our works served for our main goals. In consequence, these training students can not independently handle these projects by themselves. They are good helpers, such as database organization, truth files, running programs, turning different parameters for optimization of the algorithms. The following **Key Research Accomplishments** are well done with these students' help.

(1). Well done the basic research work on the design of a new CAPD system for detection and diagnosis of breast cancer with cellular images. It includes the following novel modules: (1) a novel adaptive fragmentary window filtering (AFWF) algorithm for circularity enhancement which is an important feature to identify normal and cancer cells, (2) modification of tree-structured nonlinear filtering (TSF), directional wavelet transform (DWT), tree-structured wavelet transform (TSWT), segmentation, feature extraction/selection, and classification modules. (3) Declustering for isolation of cell and nuclear.

(2). Well done the project of "development and optimization of CAD modules". We have developed lots of CAD modules. At first, the trainees were given the detailed process for CAD module development, then they were given a designated projects that is related to mentors' current research work, the trainees were asked to complete the project under the mentors' direction. It helped them get thorough understanding of development of CAD modules. Moreover, what they are assigned to do is closely related to the modern development of CAD modules for breast cancer detection. This helped them pursue careers in breast cancer.

(3). Well done the generation of medical image databases and truth files: The trainees were directed by mentor to collect different images and related information, convert film-screening images to digital format with digitizer, and construct medical image databases for both film-screening and direct digital mammography. The students studied the feature of mass and MCCs under the direction of mentors.

(4). Well done the evaluation of CAD modules: so many CAD systems for mass and MCCs detection have been developed. They need retrospective analysis with testing databases, which is a suitable study for the trainees.

Reportable Outcome

Proposal Created:

1. "Computerized Analysis of Cellular Features for Breast Cancer Diagnosis" submitted to DoD as a concept award (BC023811)
2. "Multimedia Telediagnosis Via Internet and Wireless Communication Networks for Breast Cancer" submitted to DoD as an idea award (BC030691)
3. "Computer-aided Pathological Diagnosis for Breast Cancer early detection" submitted to DoD as an idea award (BC030979)

Conference Abstract:

Wei Qian, Li K, Sun XJ, Song DS and Sharma A, "Computer-aided Pathological Diagnosis for Lung Cancer Cell Detection" Proceedings of 9th IASTED International Conference on Signal and Image Processing, August 13-15, 2003, Honolulu, Hawaii, USA

Wei Qian, Zhang Y, Sun XJ, Song DS and Sankar R, "A New Computer-aided Diagnosis System for Prostate Cancer Detection on TRUS imaging" Proceedings of 9th IASTED International Conference on Signal and Image Processing, August 13-15, 2003, Honolulu, Hawaii, USA

Yue Shen, Ravi Sankar, Wei Qian, Xuejun Sun, Dansheng Song, "Fuzzy Image Segmentation For Lung Nodule Detection" Proceedings of IEEE, 2003 International Conference on Acoustics, Speech, and Signal Processing (ICASSP 03), 6-10 April 2003 in Hong Kong,

Wei Qian, Kun Li K., Jin Yu and Yao Yudong, "Feature Extraction of Cancer Cells Using Wavelet Transform" 7th World Multi-Conference on SYSTEMICS, CYBERNETICS AND INFORMATICS (SCI 2003), Orlando, FL, July 27-30, 2003

Wei Qian, Xuejun Sun and Robert Clark, "Research Project Design for Summer Training Program on CAD System" Technical Program for the U.S. Army Medical Research and Materiel Command's (USAMRMC) Era of Hope 2002 DoD Breast Cancer Research Program Meeting, Orlando, Florida, September 25-28, 2002.

Conclusions

The long-term aim of this program is to encourage undergraduates pursuing the careers on breast cancer study, and to attract their interests on the development of CAD methods for diagnosis of breast cancer. The main objectives of this training program are:

1. Learning the basic principles of mammography and image processing, mastering basic methodologies for imaging breast cancer detection.
2. Stimulating trainees' interests on breast cancer study, encouraging students to pursue their academic career on the breast cancer study.

The training program is successful. The trained undergraduates are planned to be tracked their future careers to see the achievement of training program.